شش ابزار استخراج داده برای سرعت بخشیدن به کار شما

مقدمه مترجم

داده ها در عصر حاضر نیروی حیاتی و ارزشمندترین دارایی های هر سازمان را تشکیل می دهند و هر روز میزان این داده ها افزایش می یابد. در حال حاضر سازمان ها در اطلاعات غرق شده اند، در حالی که از نبود دانش رنج می برند. این امر نشانگر این مسئله است که سازمان‌ها نتوانسته اند از دانش درون داده ها به نحو مناسبی استفاده نمایند. در درون حجم عظیمی از داده‌ها، الگوها و روابط بسیاری میان پارامترهای مختلف به صورت پنهان باقی میماند که برای برنامه ریزی های استراتژیک و طولانی‌مدت میتواند حیاتی باشد؛ بنابراین نیاز به ابزاری است تا داده ها را به گونه ای پردازش کنند تا دانش حاصل از آن را در اختیار تصمیم گیران سازمان قرار دهند. در متن زیر شش ابزار معرفی شده است که میتواند در وقت و هزینه شما صرفه جویی کند ضمن اینکه اکثر آن‌ها رایگان هستند.

وظایف روزمره اداره یک سازمان می‌تواند سخت باشد، اما این به معنی وقت‌گیر بودن آن نیست. به ویژه زمانی که صحبت از داده‌های زیاد می‌شود، تنها راه دستیابی به داده‌های مفید، استخراج آن‌ها از پی. دی. اف.[1] یا صفحات گسترده، پاک کردن، سازماندهی، ذخیره و تحلیل است. با داشتن داده‌های زیاد، این کار می‌تواند به سرعت به یک کار تمام‌وقت تبدیل شود. خوشبختانه ابزارهای استخراج داده برای کمک، اینجا هستند.

اما اول:

استخراج داده چیست؟

استخراج داده‌ها عملیات خارج کردن داده از منبع داده است، به صورتی که بتوانید آن را پردازش یا ذخیره کنید. در اصل تحویل داده از یک قالب و دریافت آن در قالبی که موردنظر و قابل استفاده شما باشد. با استفاده از استخراج داده‌ها در یک کسب‌وکار کوچک به دنبال این هستیم تا بتوانیم داده‌ها را به سرعت و با دقت از فاکتورهای پی. دی. اف صورتحساب و فرم سفارش، قراردادها و اسناد منابع انسانی، توافق‌نامه‌های ضمانت‌نامه و موارد دیگر دریافت کنیم. پس از استخراج داده‌ها، می‌توان آنها را به سیستم مورد نظر خود وارد کرده و از آن استفاده کنید.

تنها یک مشکل وجود دارد:

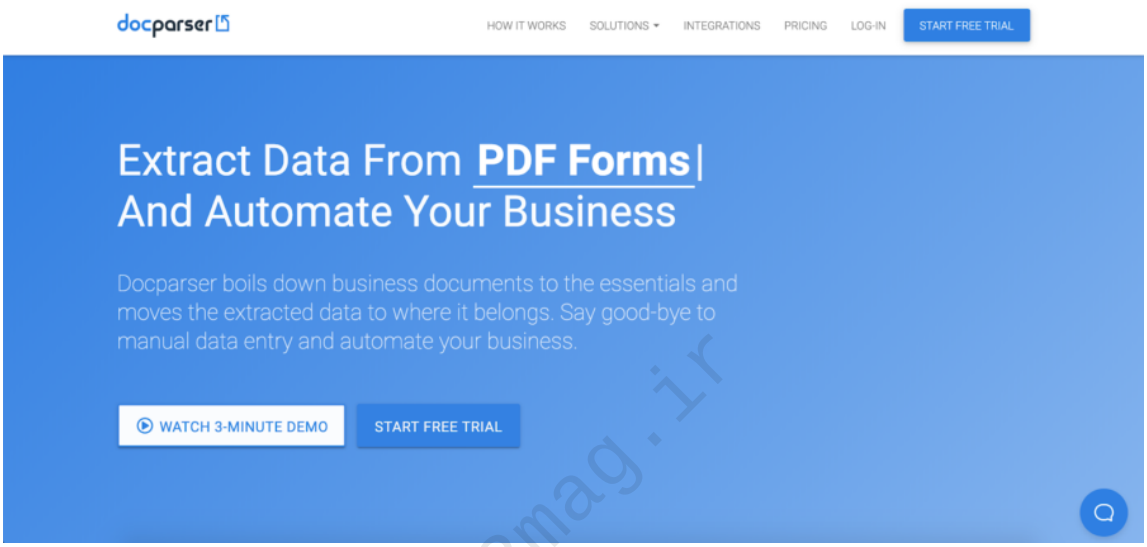
بسیاری از ابزارهای استخراج داده در دسترس هستند، بنابراین تصمیم‌گیری در مورد ابزار مناسب برای کسب‌وکار شما می‌تواند مشکل باشد؛ بنابراین چه ابزاری در دسترس است و چه چیزی برای تجارت شما مناسب است؟

شش ابزار استخراج داده

خوشبختانه فهرست برخی از بهترین ابزارهای استخراج داده موجود است. این ابزارها برای خودکارسازی فرایند مدیریت داده، صرفه‌جویی در وقت و هزینه کسب‌وکار شما وجود دارند. هر کدام ویژگی‌های مختلفی دارند، بنابراین میتوانید بررسی کرده و تصمیم بگیرید که کدام ابزارها برای کسب‌وکار شما مفیدترین هستند.

DocParser:

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.



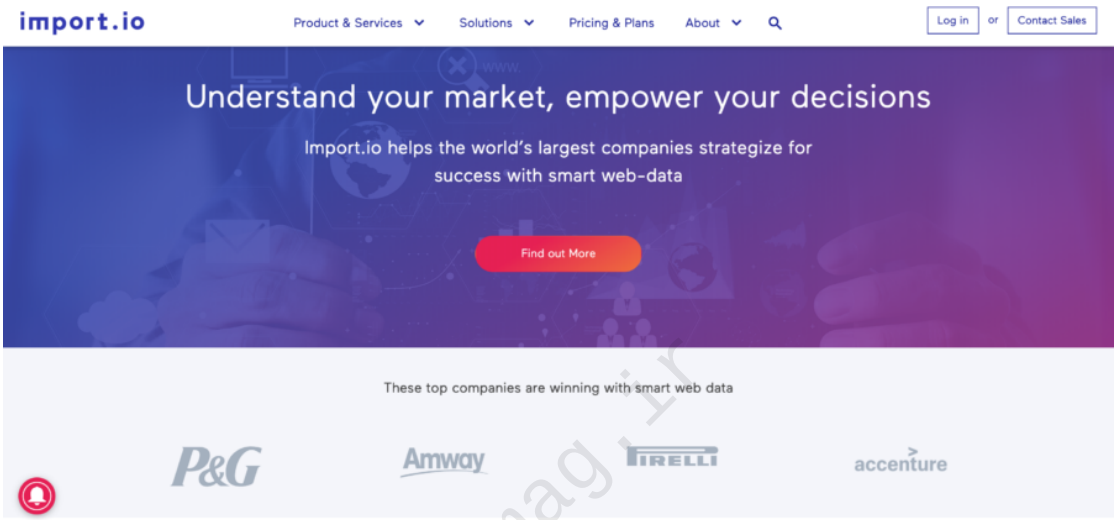
داک پارسر[۲] ابزاری با کاربرد آسان است که به شما امکان می‌دهد داده‌ها را از تمامی اسناد تجاری استخراج کنید. این برنامه یک ابزار همه‌کاره است که از یک موتور تجزیه‌کننده سفارشی استفاده می‌کند، به این معنی که می‌تواند موارد بسیاری از استفاده‌های مختلف را پشتیبانی کند. داک پارسر همچنین می‌تواند داده‌ها را از منابعی که تنها صفحات وب نیستند، استخراج کند. داک پارسر یک برنامه با صرفه‌جویی در زمان و هزینه زیاد برای بسیاری از امور در صنایع مختلف است.

موارد کاربرد این ابزار:

- فاکتورهای پی. دی. اف
- فاکتورهای اسکن شده
- سفارش‌های خرید
- سفارش‌های فروش
- قراردادها
- توافق‌نامه‌های ضمانت‌نامه
- فرم منابع انسانی (نیروی کار)
- یادداشت‌های تحویل
- سفارش‌های حمل و نقل
- فهرست‌ها
- فهرست قیمت‌ها
- اظهارات بانکی

Import.io

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.

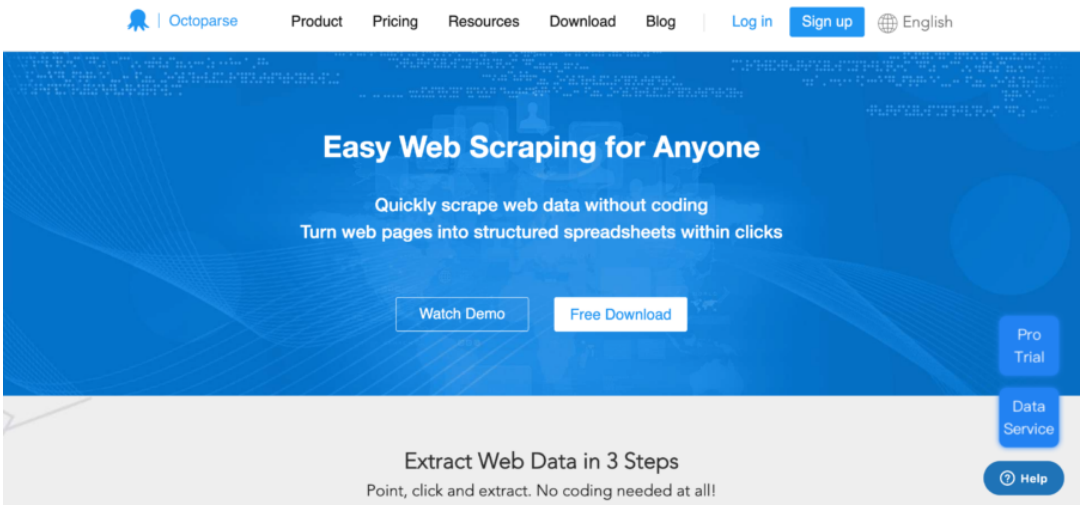


ایمپورت آیو[۳] یک ساختار یا ابزار مبتنی بر وب است که می‌تواند بدون نیاز به نوشتن کد، داده‌ها را از وب‌سایت‌ها استخراج کند. یک نسخه رایگان و همچنین یک اشتراک پولی نیز وجود دارد که خدمات مدیریت شده را ارائه می‌دهد (کارشناسان داده‌های وب از جنبه فنی مشکلات شما را پشتیبانی می‌کنند). این ویژگی یک مزیت عالی است که شما نیازی به ساختن هر چیزی از ابتدا ندارید، و این کار را برای کاربران در تمام سطوح مهارتی (و مقرون به صرفه) نسبتاً قابل دسترس می‌کند. در مقایسه با استخراج داده‌های دستی، io.Import ۸ برابر بیشتر داده و ۲۰ برابر دقت بیشتری را ارائه می‌دهد، در حالی که ۶۶% هزینه‌ها را کاهش می‌دهد. این آمار برای مشاغل کوچک، که می‌توانند از این صرفه‌جویی در هزینه‌های خود بهره‌مند شوند، مفید است.

کاربردها و قابلیت‌ها:

- پژوهش‌های منصفانه و داده‌های جایگزین
- تجارت الکترونیکی و خرده‌فروشی
- سفر آنلاین
- هوش تجاری و بازاریابی
- مدیریت ریسک

Octoparse



اکتو پارس[۴] یک فرآیند سه مرحله‌ای ساده برای جمع‌آوری داده‌ها است که باز هم نیازی به کد نویسی ندارد و فقط کافی است داده‌های مورد نظر خود را کلیک کرده و استخراج کنید. این برنامه این امکان را به شما می‌دهد تا هر وب‌سایتی را

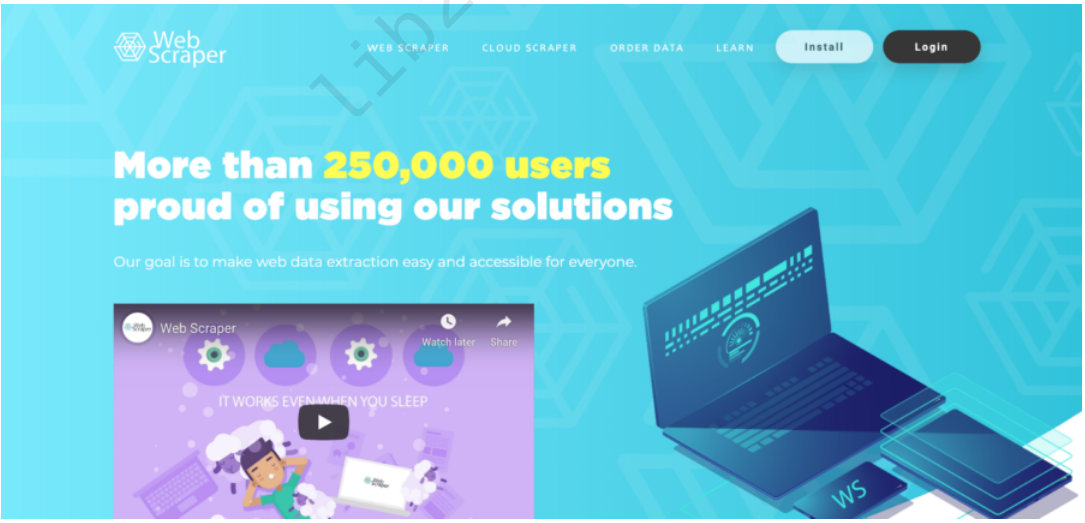
تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.

ضبط کنید، حتی آن‌هایی که از پیمایش نامحدود استفاده می‌کنند یا از ملزومات ارتباطی‌شان، ورود به سیستم است. اکتو پارس برای جلوگیری از مسدود شدن آدرس آی. پی.[۵] از چرخش آی. پی. خودکار استفاده می‌کند، بنابراین می‌توانید وب‌سایت‌های بیشتری را ضبط کنید. با کاوش برنامه‌ریزی شده و یک رابط کاربری ساده، هر فردی در تیم شما که چگونگی مرور اینترنت را می‌داند می‌تواند از این ابزار استفاده کند.

ویژگی‌های اکتوپارس:

- نظارت بر قیمت
- تولید مدرک
- بازاریابی
- پژوهش

Web Scraper



وب اسکرپر[۶] ابزاری برای استخراج داده‌های وب با اشاره و کلیک آسان است. هدف این برنامه این است که استخراج داده‌ها برای همگان آسان باشد. وب اسکرپر داده‌ها را از سایت‌هایی استخراج می‌کند که معمولاً بدست‌آوردن داده از آن‌ها سخت‌تر است، از جمله این موارد پیمایش چند سطحی، جاوا اسکریپت[۷] یا پیمایش نامحدود هستند.

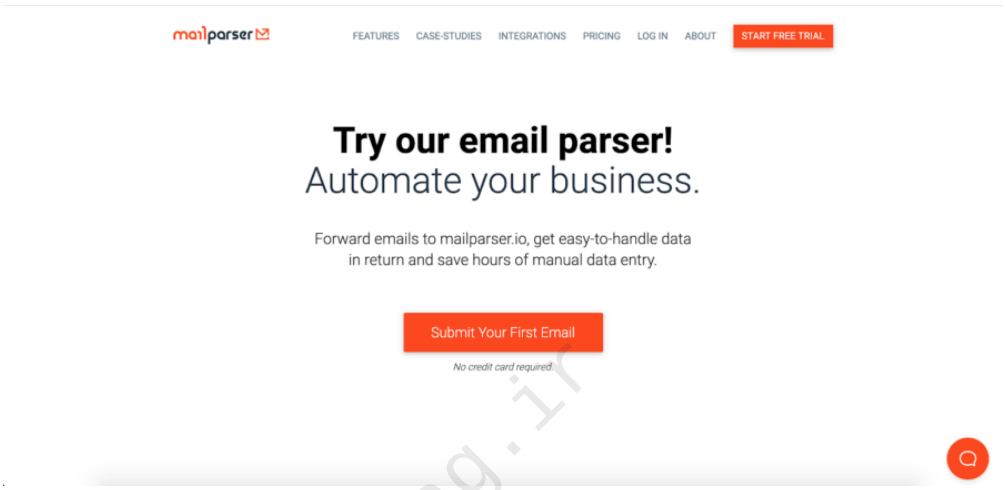
این برنامه هم به عنوان یک برنامه افزودنی مرورگر رایگان و هم به عنوان اشتراک ماهیانه، برای عملکردهای بیشتر و چندکاربره در دسترس است. یک نکته جالب در مورد این ابزار وب این است که بر روی فناوری ابری ساخته شده است. این ویژگی به شما این امکان را می‌دهد تا با شغل خود رشد کنید، بنابراین لازم نیست نگران رشد، توسعه و تغییر ابزار باشید.

قابلیت‌ها و کاربردهای این نرم‌افزار:

- کاوش سایت‌های تجارت الکترونیکی
- استخراج چندین رکورد از یک صفحه واحد
- گرفتن اطلاعات صفحه محصول

Mailparser

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.

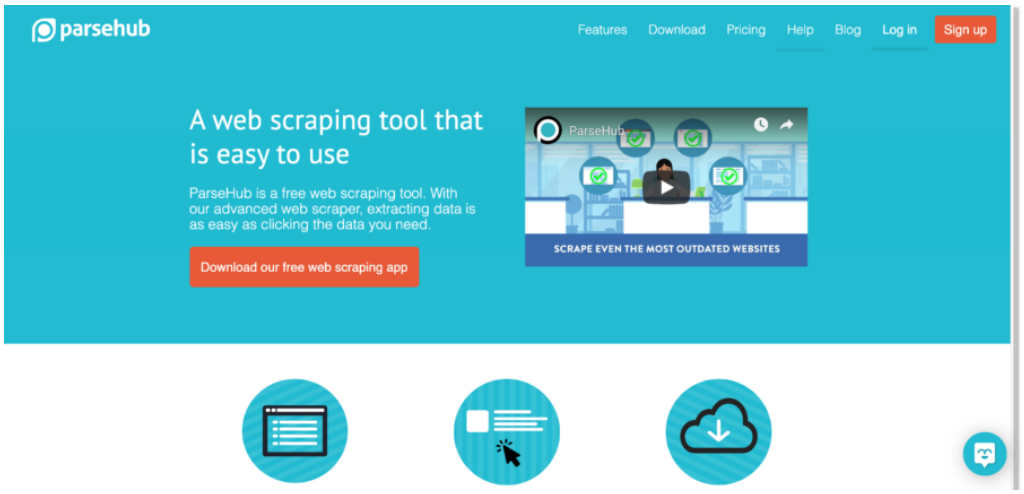


میل پارسر[۸] داده‌ها را از ایمیل استخراج می‌کند، بنابراین لازم نیست که به صورت دستی آن را وارد کنید. در نتیجه باعث صرفه‌جویی در وقت شما می شود. استفاده از آن بسیار ساده است ، فقط ایمیل‌ها را به میل پارسر هدایت کنید، تا داده‌های مورد نظر شما بر اساس قوانین استخراج سفارشی شخصی شما بیرون کشیده شود. پس از استخراج داده‌ها، می‌توانید آن را دانلود کرده و یا ادغام کنید تا در جایی که لازم باشد از آن استفاده کنید.

قابلیت‌ها و کاربردها:

- ورود اطلاعات خودکار
- پرس‌وجوهای تماس
- ضبط مدرک
- تدارکات و تحویل
- تکمیل سفارش
- تجزیه پیوست
- تجارت الکترونیک
- مشاور املاک
- گردشگری
- دیجیتال و ارتباطات
- خدمات خانگی

ParseHub



پارس هاب[۹] ابزاری برای کاویدن و استخراج داده‌ها از وب است که به شما امکان می‌دهد داده‌ها را با یک کلیک استخراج

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.

کنید. این سایت می‌تواند وب‌سایت‌های پیچیده‌ای که از جاوا اسکریپت یا آژاکس[۱۰] استفاده می‌کنند و نیز سایت‌هایی که محتوا را با ورود به سیستم محدود می‌کنند یا از پیمایش نامحدود استفاده می‌کنند را مورد استفاده قرار دهد. داده‌های استخراج شده در قالب‌های جیسون[۱۱]، اکسل[۱۲] یا ای.پی.آی. [۱۳] برگردانده می‌شوند، بنابراین می‌توان به سهولت در ساختار یا بستر مورد نظر خود استفاده کرد.

کاربردی برای:

- تحلیلگران و مشاوران
- رهبران فروش
- توسعه‌دهندگان
- مصالح تجاری و بازارها
- دانشمندان داده و روزنامه‌نگاران
- تجارت الکترونیک

خلاصه

شش ابزار استخراج اطلاعات مفید در این مطلب معرفی شد تا بتوانید از آن‌ها در کار خود استفاده کنید. با وجود بسیاری از ابزارهای ارائه‌شده خودکار، می‌توانید بسیاری از کارهای دستی وقت‌گیر شغل خود را کاهش داده، همچنین میزان خطا را به حداقل برسانید. به ویژه مشاغل کوچک می‌توانند از ابزارهای استخراج داده بهره‌مند شوند، چنان که با گروه‌های کوچک می‌توانند کارهای بزرگ‌تری انجام دهند.

ممکن است شما آگاه نباشید، اما صفحه گسترده در واقع می‌تواند یک وب‌سایت عالی باشد. در حقیقت ضبط وب با گوگل شیتز[۱۴] بسیار ساده است و حتی الگوهای رایگانی نیز در دسترس است. بنابراین قبل از اقدام به تهیه یک ابزار پولی، ببینید ابتدا می‌توانید چه ابزارهایی را به صورت رایگان در اختیار داشته باشید.

اگر در فعالیت‌های خود از داده‌ها استفاده می‌کنید، برخی از این ابزارهای استخراج داده را آزمون کنید تا ببینید که چقدر به شما کمک می‌کنند.

پانویس ها

[۱] PDF (Portable Document Format)

[۲] DocParser

[۳] Import.io

[۴] Octoparse

[۵] IP

[۶] Web Scraper

[۷] Java Script

[۸] MailParser

[۹] ParseHub

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.

Ajax [۱۰]

JSON [۱۱]

Excel [۱۲]

API [۱۳]

Google Sheets [۱۴]

منبع

این نوشته ترجمه مطلبی است با عنوان

WorkData Extraction Tools to Speed Up Your

lib2mag.ir

مشخصات استناددهی به این مقاله

نویسنده(ها):	حمیده جعفری پاورسی
عنوان مقاله:	شش ابزار استخراج داده برای سرعت بخشیدن به کار شما
عنوان مجله:	کتابدار ۲.۰ –(عنوان لاتین: i-Kitbdr 2.0)
دوره مجله(Vol):	۶
شماره مجله(Issue):	۴
سال(Year):	۱۳۹۹
شناسه دیجیتال(DOI):	
لینک کوتاه:	http://lib2mag.ir/11943

تمامی حقوق برای مجله کتابدار 2.0 محفوظ است. نقل مطالب مجله تنها با ذکر منبع و نام نویسنده مجاز است.